DOCUMENT RESUME

ED 213 752 TM 820 167

AUTHOR Martin, Paula
TITLE Evaluation in Bilingual Education.
PUB DATE Oct 81
NOTE 16p.; Paper presented at the Joint Annual Meeting of the Evaluation Network and Evaluation Research Society (Austin, TX, September 30 - October 3, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Bilingual Education; Elementary Secondary Education; *Evaluation Methods; Program Effectiveness; *Program Evaluation
IDENTIFIERS California; *Elementary Secondary Education Act Title VII; Evaluation Problems; *Meta Evaluation

ABSTRACT

The need to improve evaluation practices in Elementary Secondary Education Act (ESEA) Title VII bilingual education programs is addressed. The study examined existing efforts in bilingual program evaluations through a meta-evaluation of a sample of ESEA Title VII projects funded in 1979-80 that were representative of five regions of California. This meta-evaluation was done by examining the evaluation designs as stated in project proposals and the actual implementation of evaluation activities in the projects' evaluation reports that were submitted to the U.S. Department of Education. Evaluation designs frequently emphasized an evaluation based on instructional objectives in relation to student achievement. Evaluation reports focused on student achievement data and included survey questionnaire results regarding staff development and community involvement in some cases. Survey results from California Title VII Directors indicated that their technical assistance needs for conducting Title VII evaluations involved measuring the affective domain, using evaluation information effectively, developing in-house evaluation capabilities and analyzing data. The meta-evaluation results of this study indicated that Title VII projects still need to improve their evaluation designs and reports before evaluation reports can truly be considered as databases for future meta-analyses or as statements on the effectiveness of bilingual education. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED213752

Paper presented at the Joint Meeting
of the Evaluation Network and
Evaluation Research Society,
Austin, Texas - October, 1981

Paula Martin, Ed.D.
California Evaluation Services
2090 Green Street, No. 27
San Francisco, CA 94123
(415) 922-6528

CES 9/81

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Evaluation in Bilingual Education

I. Bilingual Evaluation Issues and Needs

There is a current need for data on the effectiveness of federally-funded bilingual education programs. ESEA Title VII serves as the major funding source for bilingual education in public schools throughout the nation. Most bilingual programs are supplemented by local and state funds, but Title VII funding is essentially the major funding base for bilingual education programs. In fiscal year 1979, $158,600,000 was spent on federally-funded bilingual education, and $166,962,500 was spent in fiscal year 1980. The Title VII budget for 1981 is $205,695,000. Congress has recently set a ceiling on Title VII funding for FYs '82, '83, and '84 at $139,000,000 for programs presently operating under Title VII. Given the current and proposed cuts in federally-funded programs, there is an immediate need to account for the cost-benefit effects of Title VII to ensure continued funding at an appropriate level that sufficiently supports implementation efforts.

Since the passage of the Bilingual Education Act in 1968, bilingual education has continually been a subject of controversy as to whether it is the best possible apporach in educating limited-English-speakers.

Evidence has shown that a great variation in the designs and reporting of bilingual program evaluations has limited the use of

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
P. H. Martin
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 820 167

© 1981 Paula Martin

evaluation results in making generalizable conclusions about program implementation and the effects of bilingual education. Information is needed on the effectiveness of various program designs, the impact of specific program designs, the impact of specific program components, and which instructional approaches are most successful.

Bilingual education programs, instructional models, and methodologies run a wide range of possibilities in program design. Evaluation strategies have to be adaptable to various program designs and must attempt to gather data regarding overall program outcomes and the programmatic variables, or specific instructional models, that caused those outcomes.

Students in bilingual education programs are impacted by other categorical aide programs like Title I and state monies. It is difficult to isolate the various effects of these instructional programs and additional support when a multiple funding/multiple treatment effect leads to many confounding variables in program treatment.

Another concern in bilingual program evaluations has been appropriate assessment instruments for limited-English-speakers. There is also a shortage of qualified evaluation personnel who are bilingual program specialists as well.

In 1982 the Office of Bilingual Education and Minority Languages Affairs (OBEMLA) in Washington, D.C. will present data to Congress on the effects of federally-funded bilingual education efforts to date for the reauthorization of ESEA Title VII. There could be more concrete data available if Title VII projects across the nation followed a more standardized reporting format and improved their data collection efforts through their annual evaluation process. Data could then be aggregated on a representative sample of projects from various parts of the country and from projects

of various sizes. An increase in evaluation commitments and interests would provide valid data on bilingual education that could be used to document program efforts and improve existing projects.

Five issues and evaluation needs are presented in Table I to address what could be done at a local, state, and federal level to improve Title VII evaluations.

There is a continued need to address the issue of a lack of national data on the effects of bilingual education programs. Meta-analysis studies are needed to aggregate data state-wide and nationally. There have been several large scale meta-analysis studies regarding the effectiveness of ESEA Title I programs for compensatory education. Data aggregation for Title VII and a national meta-analysis study have yet to be conducted. Before these types of studies can be done, there must be an improvement in general data collection through Title VII evaluations and special studies.

There are possibilities for aggregating data, especially on federally-funded projects, through the solutions outlined in Table I.

Government and state education agencies have only to be clear on their intentions for data gathering and offer federally-funded and/or state-funded projects more specifics and guidance in meeting evaluation requirements. The lack of valid and reliable data is due to an uncertainty on the behalf of local education agencies regarding the amount of funds that should be put into evaluation efforts and what data they should be collecting. Standardized evaluation formats, questionnaires, and specific directions for collecting, analyzing, and presenting student achievement data are necessary. This could be done through a handbook approach backed up with technical assistance and an evident interest in evaluations from

Evaluation Issues and Needs

TABLE I

Evaluation Issues and Needs	Solution
I. The Need for Information/Data for Various Audiences A. Local Students/Parents/Teachers/ Principals/Board Members B. State C. Federal	1. Various Reports for Different Audiences
II. The Need for Evaluation Resources A. Personnel: Evaluators, Assistants, Clerical Assistance B. Funding for the Evaluation C. Additional Consulting or Technical Assistance	1. Appropriate budgeting 2. Federal interest, encouragement, and guidelines for evaluation costs 3. Graduate students 4. Evaluation, Dissemination, and Assessment Centers (EDACs); Lau Centers; and State Departments
III. The Need to Increase Project Capability to Conduct Evaluations	1. Clarification on state and federal information needs, regulations, and guidelines 2. Technical assistance 3. Standardized reporting format, checklists, questionnaires, and scoring formats in a walk-through evaluation handbook approach
IV. The Need to Develop Interest in Evaluations	1. Incentives (positive and negative) at a state and federal level 2. Meta-evaluation feedback (state and federal) 3. Federal pulling of funds or rejection of evaluations for poor quality
V. The Need for Special Studies	1. Request for proposals for specially-funded contracts to conduct controlled evaluations on selected projects

funding agencies as reflected through meta-evaluation feedback and incentives (positive or negative) to encourage LEAs to do appropriate data collection.

II. Research and Bilingual Program Evaluation

There has been a demonstrated need in the area of bilingual education for an improvement in approach, methodologies, and data collection in program evaluation. It is evident from research (Zappert and Cruz, 1977; Alkin, 1976; California State Department of Education, 1979; Dulay and Burt, 1979; and Martin, 1981) that there is a need to improve the quality of data collected annually through Title VII program evaluations.

Overall conclusions on the lack of reliability and validity in research and evaluation studies have pointed toward the following concerns:

1. Failure to operationally define the bilingual education program--its general intentions, goals, objectives, and desired outcomes.

2. Unreliable methods are used to identify limited-English-proficient students.

3. Non-corresponding definitions for non-English and limited-English speakers across various studies.

4. Students are mobile/transient, and there may be a high attrition rate.

5. Inappropriate test selection and administration.

6. Tests are not congruent with the curriculum of the bilingual program.

7. Inadequate time allocation between pretests and posttests.

8. Improper sampling techniques / no control grouping / nonrelevant comparisons.

9. Limited-English-proficient students may not be represented in the norming sample when using standardized tests for reference points.

10. No baseline data.

11. Inappropriate data analyses / improper selection of statistical techniques.

12. No control for confounding variables:
    a. initial language dominance
    b. socioeconomic characteristics
    c. time in program
    d. balance of language use in instruction
    e. curriculum and materials in use
    f. instructional methodologies
    g. teacher and aide characteristics/qualifications
    h. other implementation characteristics.

13. Lack of adequate budgeting for evaluation efforts.

III. Evaluation Requirements for Bilingual Education Programs

Evaluation responsibilities for bilingual education involve federal guidelines, state requirements, and local needs.

Title VII requirements for bilingual program evaluations are stated in the Federal Register as the "Title VII Regulations" (Federal Register, 1980). Evaluation requirements are stated as follows:

> 123a.23
> (1) Provisions for measuring the accomplishment of the instructional objectives of the project;
> (2) Provisions for measuring the students' progress in improving their English language skills; and
> (3) A procedure for using the information gained from the evaluation to improve the operation of the project

Furthermore, in reviewing applications for Basic Projects to be funded through Title VII, evaluation plans are given 15 points within the possible 110 points for entire applications. Current regulations state:

> 123a.30
> (e) Evaluation plan (15 points).
> The Commissioner reviews the evaluation plan and its relation to the scope and objectives of the project. Specifically, the Commissioner considers the extent to which:
> (1) The overall evaluation design is consistent with, and appropriate for the insturctional and training objectives of the project;

(2) The design includes attention to:
  (i) Assessing attainment of all objectives;
  (ii) Data collection instruments and methods;
  (iii) Data analysis procedures;
  (iv) Time schedules; and
  (v) Staff responsibilities;
(3) The evaluation design specifies a comparison procedure to estimate what the performance of participants would have been in the absence of the project;
(4) The methods to be used to identify nonparticipants for comparison or to utilize another comparison standard (e.g., a historical or statistical comparison) have been identified;
(5) Sampling procedures have been identified, if applicable, and steps have been taken to ensure that the sample results will be representative of the entire project population.
(6) Data collection and analysis procedures have been identified which will address evaluation questions and which are appropriate for use with the project data.

Each State Department of Education will have its own regulations regarding the evaluation of bilingual education. These regulations will be tied into state legislation for bilingual education. There may be program reviews by the state department to monitor and evaluate state-funded bilingual efforts.

IV. Characteristics of California Bilingual Program Evaluations: FY 1979-80

In FY 1978 California received 35.4% of all available Title VII funds. There were 146 Title VII projects, with New York having the second largest number of projects (92) and Texas having 70 projects. In 1980-81 California had 128 Title VII projects.

The author's dissertation study involved a meta-evaluation of 43 evaluation designs and 47 evaluation reports of California Title VII projects that were funded in FY 1979-80.

A review of 43 evaluation designs indicated that the evaluation approach most frequently proposed was based on instructional objectives

in relation to student achievement with an emphasis on the evaluation of the instructional component. Data collection procedures proposed questionnaires to staff, parents, and community. Interview techniques and site-observations were also proposed, but less frequently. A review of the evaluation designs indicated that there was a need for more coverage on data collection techniques, timelines for evaluation activities and testing, a description of the tests and instruments to be utilized, and definition of evaluation resources in terms of funds and staffing. Analytic procedures for student achievement data were stated in very general terms, with the most commonly proposed approaches employing means, percentiles, grade-equivalent scores, standard deviations, or t-tests.

Evaluation reports focused on student achievement data and included survey questionnaire results in some cases. Three standardized instruments were commonly used for student achievement data in English. One-half of the 47 projects also measured student achievement in Spanish, using several standardized instruments. There were three commonly used language dominance instruments and continuums of skills were used in some cases for English as a Second Language (ESL) assessment. Criterion-referenced tests, although mentioned frequently in evaluation designs, were not mentioned to any extent in evaluation reports.

The most common analytic techniques involved pre/post testing on a fall-spring basis, the use of raw scores, percentiles, and grade-equivalent scores. Projects had a tendency to report mean or median raw scores, percentiles, or grade-equivalent scores. Scores were frequently reported by grade only and not by language proficiency, so non-English-speakers' scores were often analyzed with fluent-English-speakers' scores. The N (number) of students was not included in averaged student achievement

data in many cases. Correlations, Chi-square, and analyses of variance were seldom used. A few projects analyzed test scores in relation to student attendance. A longitudinal approach to student achievement was alluded to, but only two projects fully implemented this design. Sampling techniques for any purposes were not covered in sufficient detail, and control/comparison groups were not utilized to any extent or described. In general, there was not a sufficient interpretation of student achievement data or overall evaluation results.

Evaluation utility was apparent in only half of the projects for program planning and reports to several audiences. Conclusions and recommendations could have been stronger in many cases.

Evaluation resources were more apparent in evaluation reports: 11 of the 47 projects contracted with medium-sized research firms to do their evaluations, 17 went to external evaluation consultants, twelve were internal evaluations, and 2 evaluations were done by a consortium and a university-based research service. Staffing for the evaluation was not evident in 5 reports.

Results of this study indicate that Title VII evaluation efforts still need to be planned in more detail, with additional technical assistance regarding desired outcomes through data collection and analyses of student achievement scores. Ongoing or formative evaluation efforts need to be more apparent. Projects also need more ideas regarding successful questionnaire layouts and ways to display student achievement data. Reports need to include a more in-depth analysis of test scores, stronger conclusions, and specific recommendations that can actually be used in future program planning and revisions.

V. Results of a Bilingual Evaluation Technical Assistance Needs Questionnaire to Title VII Project Directors (1980)

In May and June 1980 a questionnaire was sent to 128 Title VII bilingual project directors in California to determine their needs for technical assistance in planning and conducting Title VII bilingual program evaluations as another aspect of the author's dissertation study. There were ninety-seven returns on the survey, indicating a 76% response rate, with four districts choosing not to participate in the study. Results of the survey were then representative of ninety-three projects, or 73% of the total population. Respondents included eighty-three project directors, five program evaluators, and five "other" personnel involved with the Title VII projects.

Results indicated that districts are most interested in how evaluation information can be used more effectively and have concerns regarding the development of in-house evaluation capabilities. Analyzing evaluation data and student achievement in the second/native language were also considered to be areas where technical assistance was most needed.

Areas of secondary evaluation needs were the development of specific instruments and questionnaires, interpretation of the results of data analyses, and the preparation and writing of reports. The development of the overall evaluation design, the selection of appropriate tests, and general measurement of student achievement were lesser concerns. Least problematic were data collection procedures, English language assessment, and hiring an external evaluator.

There was an interrelationship of needs concerning major evaluation components such as design development, student achievement data (both languages), the selection of appropriate tests, and especially the development or selection of specific instruments and questionnaires.

VI. The Multiple Funding Issue and Bilingual Program Evaluations

There are given challenges in approaching an evaluation of a bilingual program/project given the implications of the multiple funding / multiple treatment issue.

In the evaluation of a bilingual program, consideration must be given to evaluation needs at several levels and meeting evaluation requirements of several funding sources. Most bilingual programs operate through local, state, and federal funds and they must consider the collection of evaluation data for accountability purposes for several audiences.

Compensatory education programs cover a wide range of students with special needs. Many students are impacted by several categorical aid programs. Due to this overlap of services, the multiple funding / multiple treatment issue needs to be closely examined if one is to have a valid program evaluation concerning any one treatment.

Practical concerns regarding collecting student achievement data when students are in multiple-funded programs and schools can be a complicated issue. Out-of-level testing, testing procedures, and instruments have to be considered within the framework of whatever testing may be taking place by the school district in terms of minimum standards and Title I. Since Title I students are students that score below the fiftieth percentile, any Title I testing may affect those limited English proficient students who are Title I but also are in a bilingual program.

This points to a need for better methodologies in the collection of achievement data for students in bilingual programs and better planning for a valid program evaluation that can make generalizations regarding student gains due to programmatic variables. Various options on evaluation designs have to be explored. Any proposed evaluation design has

to take this into consideration and work within a district-wide evaluation plan and testing framework.

References

Abert, J., & Kramrass, M. (Eds.). Social experiments and social program evaluation. Cambridge, Mass.: Ballinger Publishing Co., 1974.

Alkin, M., Kosecoff, J., Fitz-Gibbon, C., & Seligman, R. Evaluation and decision making: The Title VII experience. CSE Monograph Service in Evaluation, No. 4. Los Angeles: University of California, Center for the Study of Evaluation, 1976.

Ball, Samuel (Ed.). Assessing and interpreting outcomes. San Francisco: Jossey-Bass Publishers, No. 9, 1981.

Berke, Iris Polk. Evaluation Into Policy: Bilingual Education, 1978. Unpublished dissertation. Stanford University: Stanford, CA, 1980.

Bilingual needs report 1977-78. Sacramento: California State Department of Education.

Bilingual program, policy and assessment issues. Sacramento: California State Department of Education, 1980.

Boruch, R. F. (Ed.) Secondary analysis. San Francisco: Jossey-Bass Publishers, 1978.

Bryant, F. B. & Wortmen, P. M. Secondary analysis: The case for data archives. American Psychologist, April 1978.

Burstein, L. Secondary analysis: An important resource for educational research and evaluation. Educational Researcher, May 1978.

Burt and Dulay. Research on the effectiveness of bilingual education. Educational Evaluation and Policy Analysis, 1979.

Development of evaluation and data gathering models for ESEA Title VII bilingual education projects, August 1980-April 1982. Rosslyn, Va.: Inter-American Research Corporation.

ESEA, Title I: A reanalysis and synthesis of evaluation data from fiscal year 1965 through 1970. (AIR-27400-3/72). Palo Alto: American Institute for Research.

Evaluation of the classroom instruction component of the ESEA Title VII Bilingual Education Program. Development Associates, Inc. in affiliation with Abt Associates, Inc. for the Office of Program Evaluation, U.S. Department of Education.

Evaluation of the impact of ESEA Title VII Spanish/English bilingual education programs: Overview of study and findings. (AIR-48300-3/78-FR VI). Palo Alto: American Institute for Research, 1978. (ERIC Report ED 154-6340).

Evaluation report of multiple-funded programs, 1976-77. Sacramento: California State Department of Education, 1978.

Federal Register, April 1980, Part 3, p. 23216.

Forum, December 1980, April 1981, September 1981

Glass, G. V. Integrating findings: The meta-analysis of research. Review of Research in Education, 1977, pp. 351-379.

_____. Policy for the unpredictable (uncertain research and policy). Educational Researcher, October 1979.

_____. Primary, secondary and meta-analysis of research. Educational Researcher, December 1976, pp. 3-8.

_____. Standards and criteria. Occasional Paper Series, No. 10. Kalamazoo, Mich.: Western Michigan University, The Evaluation Center, 1977.

_____, McGaw, B., & Smith, M. L. Meta-analysis in social research. Beverly Hills, Calif.: Sage Publications, 1981.

Horst, D. P., Tallmadge, G. K., & Wood, C. T. A practical guide to measuring project impact on student achievement. No. 1 in a series of monographs on evaluation in education. Washington, D.C.: Government Printing Office, No. 017-080-0146-2. Prepared for HEW by RMC Research Corporation, 1975.

Law, A. Evaluating bilingual programs. TM Report 61. Princeton, N.J.: Educational Testing Service, ERIC Clearinghouse on Tests, Measurement, and Evaluation.

Law, A. Handbook for the evaluation of multi-funded educational programs. (2nd rev. ed.). Sacramento: California State Department of Education, Office of Program Evaluation and Research, 1975.

Light, R. J. Capitalizing on variation: How conflicting research findings can be helpful for policy. Educational Researcher, 1979, 9(9).

_____. Research integration: The state of the art. Evaluation in Education, 1980, 4(1).

_____, & Smith, P. V. Accumulating evidence: Procedures for resolving contradictions among different studies. Harvard Educational Review, 1971, 41, 429-471.

Locks, N., & Pletcher, B. A. Language assessment instruments for limited-English-speaking students: A needs analysis. U.S. Department of Health, Education, and Welfare, National Institute of Education, 1978.

Longitudinal Study Title VII Bilingual Program, Santa Fe Public Schools, Santa Fe, New Mexico. Los Angeles: California State University, The National Dissemination and Assessment Center, 1978.

Lopez, G., & Cervantes, R. Abstract of AIR report criticism. Sacramento: California State Department of Education, Bilingual-Bicultural Section, 1978.

Pillemer, D. B., & Light, R. J. Using the results of controlled experiments to construct social programs: Three caveats. Evaluation studies review annual. Beverly Hills, Calif.: Sage Publishing Company, 1979

Rentz, C. C., & Rentz, R. R. (Eds.). Evaluating federally sponsored programs. San Francisco: Jossey-Bass Publishers, No. 2, 1978.

Roberts, K., & Burstein, L. (Eds.). Issues in Data Aggregation, No. 6. San Francisco: Jossey-Bass Publishers, 1980.

Sanders, J. R., & Nafziger, D. H. A basis for determining the adequacy of evaluation designs. Occasional Paper Series, No. 6. Kalamazoo, Mich.: Western Michigan University, The Evaluation Center, 1976.

Scriven, M. An introduction to meta-evaluation. In P. A. Taylor & D. M. Cowley, (Eds.), Readings in curriculum evaluation. Dubuque, Iowa: Wm. C. Brown Co. Publishers, 1972. (a)

_____. Key evaluation checklist. Pt. Reyes, Calif.: Edgepress, 1974.

_____. The methodology of evaluation. In P. A. Taylor & D. M. Cowley (Eds.), Readings in curriculum evaluation. Dubuque, Iowa: Wm. C. Brown Co. Publishers, 1972. (b)

Sheppard, Lorrie. A checklist for evaluating large scale assessment programs. Occasional Paper Series, No. 9. Kalamazoo, Mich.: Western Michigan University.

Smith, N. (Ed.). Federal efforts for developing new evaluative methods. San Francisco: Jossey-Bass Publishers, 1981.

Summary of a study of ESEA Title VII proposals, 1975-1978. Sacramento: California State Department of Education, Office of Bilingual-Bicultural Education.

Tallmadge, G. K., & Horst, D. P. A procedural guide for validation of achievement gains in educational projects. No. 2 in a series of monographs on evaluation in education. Washington, D.C.: Government Printing Office, No. 017-080-01516. Prepared for HEW by RMC Research Corporation, 1971.

Weiss, C. Using social research in public policy making. Lexington, Mass.: Lexington Books, 1977.

Zappert, L.T. & Cruz, B.R. Bilingual education: An appraisal of empirical research. BABEL - Bay Area Bilingual Education League/Lau Center, Berkeley Unified School District, 1977.